

# Re: word webpages

---

*Source:*

<http://www.tech-archive.net/Archive/Mac/microsoft.public.mac.office.word/2006-11/msg00340.html>

---

- *From:* "John McGhie [MVP – Word and Word Macintosh]" <[john@xxxxxxxxxxx](mailto:john@xxxxxxxxxxx)>
  - *Date:* Wed, 15 Nov 2006 21:59:23 +1100
- 

Hi Phillip:

My experience with DreamWeaver is very, very small. I use FrontPage extensively (because I'm webmaster for [www.word.mvps.org](http://www.word.mvps.org), and it's built with it). Oh, and I get it for free... :-)

But I'll give you a "discussion" on "Fix" as it applies to Word HTML.

The whole subject of Markup Languages is very poorly understood by the non-professional audience, and nearly all of them have completely missed the point of what Microsoft was trying to do (and has succeeded in doing...)

To begin with, Word DOES NOT write HTML. It never did, and was never designed to. Word writes XML and XHTML. When Microsoft first included the XML converter in Office, the Marketing Department came to the conclusion that all users were too stupid to understand XML, so they forced the designers to call it "Web Page" and "HTML".

It's not, and it was never designed to be :-)

The reason Microsoft chose to output markup language in XML is very simple: HTML simply isn't powerful enough to describe a Word document.

The design goal was to express a Word document in language a browser can read. The ultimate goal was to render that document so exactly that if you printed from Word and printed from the browser, you would not be able to see the difference. They wanted the user to be able to round-trip their document onto the web, then back to a document on another computer. You cannot DO that in HTML. It's not powerful enough.

Well, the reality is not quite that good, but it gets very close. If you choose Save As from Word 2004 and check the default "Save entire file into HTML" you will get an almost exact rendition of the Word document.

You will indeed have a lot of markup language in the file that most browsers cannot interpret. That's not necessarily a bad thing. It's not against the rules :-) The whole "point" of markup languages is that the file should contain information "marked up" so the display device can pick and choose

## Re: word webpages

how much of the code it wants to use. That's the way it's MEANT to work. A teleprinter or a mobile phone may discard most of the markup and just print the text. A typesetting machine will use it all and add some of its own...

If you want to remove the "rich encoding", when you do the Save As, check the button that says "Save only display information". That will get you the 99 per cent fat-free version of the code. If you copy and paste, you will get the "Full Fat" version, because Word can't tell when you copy what you are going to paste into.

One thing you will notice is that the earlier versions of Word make very liberal use of "SPAN" tags. That's because Microsoft knows that the vast majority of Word users don't use Styles, they use direct formatting. Using SPAN tags you can directly replicate the ransom-note formatting of the average school project :-). Word 2003 and later are much more likely to "coerce" direct formatting into styles, and thus to express the formatting as attributes of the <p> tag.

If you do save in the "Lite" version, don't expect to be able to re-create the Word document from the web page. The filter removes the complex information needed to rebuild the document from the web page, so you have a one-way trip.

The difference is like a 16-track studio recording compared to an MP3. The MP3 is much smaller -- but only because it has removed as much sound as it thinks you can get away with. If you have good ears (and I suspect you have...) and you can still find someone with a high-end vinyl record player, have a listen. Try to do it on a Deutsche Gramophon live recording, and compare it with the CD of the same performance. If you know what you're listening for, I bet you find that some of the magic is gone :-)

Now let's discuss "Unnecessary" as applied to web coding. First off, who says it's unnecessary? The web designer? The web coder? The reader?

And WHY would any of them decide that code is unnecessary? When I first started web coding, a faaaaast connection ran at 2,400 bps and nobody could afford 28.8 :-). Then, you kept your code damned simple if you wanted your readers to be able to access it at all.

These days, most of the material I use Word to encode is reference manuals or Help Files that are being accessed at gigabit speeds. On a gigabit LAN, even the "Full Fat" flavour of a Word document will get to the computer faster than the computer can draw the screen.

The user simply can't detect the difference between a tightly hand-coded web page using all the economy tricks and a full-featured page pasted from Word. But the customer can!! Let's assume my customer is paying \$1,000 a day for casual web coding (that's about the going rate...) and his on-costs are double the web coder's wages (and that's about normal) the customer is spending \$2,000 a day to have his material put on the web.

## Re: word webpages

For that, I can give him more than a thousand pages saved straight out of Word, or maybe 1 to 10 pages carefully crafted in DreamWeaver. About a buck a page, or between \$100 and \$2,000 a page. You do the math...

Don't forget the time involved: I'm currently working for an ISP. We just spent around \$3.6 million to get maybe 50 pages produced by a web coding company. It took about six months to get the result. We think we got good value, and we'll use them again. These were intensely complex "dynamic shopping cart" pages, with which we will make squillions, so the investment of both time and effort was worth it. And each of those pages contains several times as much code as Word produces :-)

The projects I use Word for are technical manuals that update every few months. One project I work on produces a new version of two 500-page-ish manuals each month. They make it available to their customers world-wide as HTML, which I fire straight out of Word in under a day. Yes, they could save the manuals to PDF and produce them even quicker: but most customers balk at downloading a 12-15 MB PDF, only to find it has very limited cross-referencing, searching, and interlinking.

For that sort of job, Word is your weapon. It all comes down to the business drivers you need to achieve. Word is the most amazing web page sausage-machine. If you prefer Chateaubriand, call a professional chef and his copy of DreamWeaver :-)

Yes, I can do this straight out of Word to the web server. And I can do almost all of it with Word 2004. Sadly, they left a few tricks out of Mac Word that come in handy when you're doing this sort of thing. PC Word has a tuneable output filter that enables you to determine with a high degree of control which tags will be removed when you filter, scalable vector graphics, and the ability to attach cascading style sheets. I expect these abilities will arrive in the next version of Mac Word (together with much better XML graphics handling).

My "encouragement" to Microsoft is "don't forget the Professional User!" You do have some professional users of Microsoft Word. And you're more likely to find them using a Macintosh than a PC.

Currently, the Bean Counters and their allies in the Marketing Department have decreed that Microsoft Word be massively dumbed-down in the next version. That's because they have gigabytes (literally) of research data that prove that most Word users never get beyond the notion of a "glass typewriter". So why spend the money on all these expensive powerful features that "users" never use?

Well, they do on a Mac. Maybe they don't need all those bells and whistles on the PC. But here, we do. Here, people who pay money for Mac Office do so because they need it more powerful than the alternatives, not because they want it simpler :-)

Cheers

Re: word webpages

Re: word webpages

On 14/11/06 10:34 AM, in article uVYIWx3BHHA.4992@xxxxxxxxxxxxxxxxxxxxxxxx, "Phillip Jones" <pjones1@xxxxxxxxxxxx> wrote:

John have you ever opened DreamWeaver and took it for a spin?

One of the things That is on one of the menus is "fix Microsoft HTML code"

I tried creating pages with word, I saved as HTML and opened it in a W3C Standards compatible browser such a FireFox, Mozilla. You never seen the like of strange colored tags inserted at strange places.

After I first bought Dreamweaver 7 I noticed this little item.

After detecting everything it dumped a bunch of necessary code and tags. And the document then worked as it should.

IF you can make it work straight fro creation to upload. Your far smarter that I think you are and I think you border genius now. ;-)

John McGhie [MVP – Word and Word Macintosh] wrote:

Hi Daniel:

You may be surprised. I think you will get both sound, animation, and video in web pages from the next version of Word.

Dreamweaver is a great web editor for making single pages. Word is your weapon when you need to make web pages at the rate of three or four hundred an hour, such as for large-scale reference publishing.

I have a workflow using Word that publishes a 560-page book as a website: I can do two a day including the TOC and navbars :-)

Cheers

On 13/11/06 6:31 AM, in article 38DE89F6-7598-442F-AB76-EC4954D38416@xxxxxxxxxxxxxxxx, "DanielWalters6" <DanielWalters6@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx> wrote:

Microsoft themselves are probably not going to be much more help.

If you can afford it (no disrespect) use Dreamweaver 8 – it's brill!

It's got everything you can do in word, with the added FTP

Re: word webpages

uploading etc.

--

Please reply to the newsgroup to maintain the thread. Please do not email me unless I ask you to.

John McGhie <john@xxxxxxxxxxxx>

Microsoft MVP, Word and Word for Macintosh. Business Analyst, Consultant  
Technical Writer.

Sydney, Australia +61 (0) 4 1209 1410

.